03-1040
1496.00344

HIGH QUALITY, LOW MEMORY BANDWIDTH MOTION ESTIMATION PROCESSOR

This application claims the benefit of U.S. Provisional Application No. 60/487,643, filed July 15, 2003, which is hereby incorporated by reference in its entirety.

This application is related to co-pending applications Serial No. 10/196,731, filed July 16, 2002, Serial No. 10/669,930, filed September 24, 2003, Serial No. 10/682,631, filed October 9, 2003 and Serial No. 10/___,___ (Attorney Reference no. 03-00986 / 1496.00339), filed October 22, 2003, which are hereby incorporated by reference in their entirety.

Field of the Invention

The present invention relates to a video motion estimation generally and, more particularly, to a high quality, low memory bandwidth motion estimation processor.

Background of the Invention

Designing a motion estimation approach for a very large scale integration implementation that yields high picture quality

(i.e., DVD quality) while consuming very low external memory bandwidth poses several challenges. Using a large search area for the motion estimation yields high picture quality but uses very high external memory bandwidth and large internal buffers. Using a small search area for the motion estimation results in reduced external memory bandwidth, but produces additional controls, buffering and yields low picture quality, especially in the presence of fast motion. To counter the fast motion, each target macroblock can be searched in a small, localized area. However, little or no overlap exists between localized search areas for each target macroblock, or group of macroblocks, in a worst case scenario. Thus, very high external memory bandwidth is still consumed reading reference data for each small search areas.

A first approach for motion estimation is to encode with a regular search method using small motion estimation search areas. However, the first approach suffers from a picture quality loss. A second approach is to encode with the regular search method using large motion estimation search areas. Consequences for using the large search areas include high external memory bandwidth, a large internal memory buffer and large computational complexity. A third approach is to encode with "fast" search methods that use fewer

data points (i.e., fewer calculations per search location and/or fewer search locations) and small motion estimation search areas. The fewer data points result in a moderate external memory bandwidth usage but produce a picture quality loss while adding complexity in the forms of additional control and buffering. A fourth approach is to encode with "fast" search methods and large motion estimation search area. The large search areas consume a high external memory bandwidth, large internal memory buffers and produce additional control and buffering.

Summary of the Invention

The present invention concerns an apparatus for motion estimation generally comprising a memory and a circuit. The circuit may be configured to (i) search for a first motion vector for a first current block among a plurality of first reference samples, (ii) copy a plurality of second reference samples from the memory and (iii) search for a second motion vector for a second current block among the second reference samples copied from the memory and at least a portion of the first reference samples.

The objects, features and advantages of the present invention include providing an apparatus and/or method for high

quality, low memory bandwidth motion estimation that may (i) minimize search memory size, (ii) minimize external memory bandwidth consumption, (iii) maintains a regular data from the external memory, (iv) utilize relatively low complexity control hardware, (v) maintain a reasonably large effective search area and/or (vi) ensure high picture quality.

Brief Description of the Drawings

These and other objects, features and advantages of the present invention will be apparent from the following detailed description and the appended claims and drawings in which:

FIG. 1 is a block diagram relating example current macroblocks to reference search windows;

FIG. 2 is a block diagram of an example utilization of a search memory;

FIG. 3 is a flow diagram of an example method for motion estimation;

FIG. 4 is a partial block diagram of a first example implementation of an apparatus in accordance with a preferred embodiment of the present invention; and

FIG. 5 is a partial block diagram of a second example implementation of an apparatus

Detailed Description of the Preferred Embodiments

The present invention may realize the benefits of a large search area without using high external memory bandwidth by using two small independent "boxcar" search windows. Each search window size may be set to +/-40, 32, 24 or 16 horizontal (H) integer pels by +/-24,16 or 8 vertical (V) integer pels. Each search window may be placed within a reference frame independently of each other (e.g., each with an individual (x, y) offset). When no large motion is detected in a sequence of video frames, the search windows may be placed adjacent or adjoining each other horizontally or vertically. When fast motion is detected, one of the windows may be placed near a center of a target current macroblock to capture static background and small motion while the other search window may be placed at an appropriate position to cover the large motion. The positioning, or offsets, are generally determined once per frame. Other update rates for the offsets may be implemented to meet the criteria of a particular application. Many methods may be used to determine suitable offsets, including analyzing a

history of motion vectors and/or analyzing a results of a sub-sampled search on the current frame. A number of different techniques may be used for the analysis (e.g., histograms, clustering, etc.).

Referring to FIG. 1, a block diagram 100 relating example current macroblocks to reference search windows is shown. A motion estimation may be performed by comparing blocks of current samples in a current frame 102 against reference samples in a reference frame 104. Search areas in the reference frame 104 may be determined by a position of each current block relative to a reference corner (e.g., upper left corner) of the current frame 102 and a global offset (e.g., (gx, gy)) between the current frame 102 and the reference frame 104. Details for how to determine one or more global offsets (gx, gy) may be found in the co-pending United States application Serial No. 10/196,731, filed July 16, 2002, hereby incorporated by reference in its entirety.

Use of boxcar search windows to determine a search area may be illustrated by way of example as follows. A first current block (e.g., MBn) may be located at a location or position (e.g., (x, y)) relative to the reference corner of the current frame 102. An adjoining or second current block (e.g., MBn+1) may be located

immediately to the right of the first current block MBn. A first search area or window 106 for the first current block MBn in the reference frame 104 may be calculated by adding the position vector (x, y) to the global offset vector (gx, gy) (e.g., (x+gx, y+gy)). Since the second current block MBn+1 is next to the first current block MBn in the current frame 102, a good probability generally exists that a best match for the second current block MBn+1 may be found in a second search area or window 108 offset slightly to the right from the first search window 106. Generally, the offset may be a horizontal distance equal to a horizontal size of the current blocks (e.g., 16 pels). Other offsets (e.g., horizontal and/or vertical) of the second search window 108 from the first search window 106 may be implemented to meet the criteria of a particular application.

Referring to FIG. 2, a block diagram of an example utilization of a search memory 120 is shown. Keeping the global offset (gx, gy) constant during multiple searches throughout the current frame 102 generally allows exploitation of (i) a small memory 120 (or internal buffer) for motion estimation circuitry, (ii) a maximization of data reuse for reference samples already copied into the search memory 120 and (iii) a minimization of

external memory bandwidth to copy the reference samples from an external memory. The search memory 120 may contain storage for an array of reference blocks (e.g., 6Hx3V to 8Hx3V reference blocks) or an array of reference samples (e.g., 96Hx48V to 128Hx48V pels at 16Hx16V pels per block) over which a search may be performed.

The search memory 120 may be logically partitioned to allow different phases for multiple searches to be performed substantially simultaneously. For example, the first search window 106 for the first block MBn may be stored as an array (e.g., 5Hx3V) of reference blocks, as illustrated by area 122. A macroblock column 124 (e.g., 1Hx3V to 1Hx1V array) of reference blocks may be copied from the external memory to load data from the second search window 108 for the second current block MBn+1 while the first current macroblock MBn is searched using the area 122 of the search window data.

As the motion vector search for the first current block MBn is ending, another macroblock column 128 may start to load in the search memory 120 overwriting the reference samples in the area 122 no longer useful to the search. After the first search has completed, a new motion vector search for the second current block MBn+1 may begin using the reference data stored in an area 126.

The area 126 may contain the reference samples from the second search window 108. The area 126 generally comprises a portion (e.g., 4Hx3V blocks) of the reference samples common to the first search window 106 and the second search window 108 plus the macroblock column 124 of reference samples within the second search window 108 but outside the first search window 106. In general, each successive search area for neighboring macroblocks may overlap the previous search areas. Each new target current macroblock may be searched by loading only a single new macroblock column (e.g., 16Hx48/32/16V pels) to the search window 120 from the reference frame 104. Loading a single macroblock column instead of an entire search window generally helps to minimize the external memory bandwidth consumed by the motion estimation process.

Referring to FIG. 3, a flow diagram of an example method 140 for motion estimation is shown. The method (or process) 140 generally includes determining a first global offset (e.g., gx1, gy1) and a second global offset (e.g., gx2, gy2)(e.g., block 142). The first global offset (gx1, gy1) may be applied (e.g., block 144). The first search window 108 within the reference frame 104 may then be copied from the external memory to the search memory 120 using the first global offset (gx1, gy1) (e.g., 146). The

motion estimation circuitry may then begin searching for a first motion vector for the first current block MBn (e.g., block 148). While the first search is being performed, the new macroblock column 124 within the reference frame 104 may be coped into the search memory 120 substantially simultaneously (e.g., block 150). A check may be made (e.g., decision block 152) to determine if any additional current blocks have not been searched. If the second current block MBn+1, or any other current block, has not been searched (e.g., the YES branch from decision block 152), the process may return to the search task 148 and begin the copy task 150 to load a next macroblock column from the reference frame 104, if any.

Once motion vectors have been determined for all of the current blocks in the current frame 102 (e.g., the NO branch of decision block 152), a check may be performed for additional global offsets (e.g., decision block 154). If additional global offsets (e.g., (gx2, gy2)) exist (e.g., the YES branch of decision block 154), another search of the current blocks in the current frame 102 may be performed with the new global offset (e.g., block 156). After all of the global offsets have been examined (e.g., the NO branch of decision block 156), the search process 140 may be ended.

Referring to FIG. 4, a partial block diagram of an example implementation of an apparatus (or system) 160 is shown in accordance with a preferred embodiment of the present invention. The apparatus 160 generally comprises a memory circuit 162 and a circuit (or module) 164. The memory circuit 162 may have an interface 166 coupled to an interface 168 of the circuit 164. The circuit 164 may receive framing signals (e.g., FRAME_SIG) at an interface 170. The circuit 164 may present multiple signals (e.g., MV, SAD and MODE) at an interface 172.

The circuit 164 may be implemented as a motion estimation (ME) processor circuit. The ME processor circuit 164 is generally fabricated as a single integrated circuit. The memory circuit 162 may be fabricated on another integrated circuit independently from the ME processor circuit 164. As such, the memory circuit 162 may be referred to as an external memory circuit.

The signal FRAME_SIG generally comprises multiple signals used to determine a start of a current frame for which the motion estimation is to be performed. The signal FRAME_SIG may include, but is not limited to, horizontal synchronization information, vertical synchronization information, a frame height, a frame width and a macroblock start indication. Other information may be

included in the signal FRAME_SIG to meet the criteria of a particular application.

The signal MV may define one or more motion vectors for the current block (e.g., MBn, MBn+1, etc.) with respect to a reference frame or frames at an integer-pel resolution. The number of motion vectors within the signal MV may be defined by a particular partition mode determined to be a best mode for the current block. For example, if the apparatus 100 determines that a MODE0 should be used for motion estimation, the signal MV may convey a single motion vector for the current block. If the apparatus 100 determines that a MODE3 should be used for motion estimation, the signal MV may convey four motion vectors, one for each of the four sub-blocks within the current block.

The signal SAD may define a sum of absolute difference value for the best mode determined by the apparatus 100. The signal SAD may include a bias value and/or a cost value incorporated before determining the best mode. The bias value and the cost value may be determined from quantization parameters and the motion vectors.

The signal MODE may define the particular partition mode resulting in the best motion estimation for the current block. The

signal MODE may identify seven different modes in designs of the apparatus 160 that may be capable of partitioning each of the current blocks into sixteen smallest sub-blocks. The signal MODE may identify four different modes in designs of the apparatus 160 that may be capable of partitioning each of the current blocks into four 8x8 sub-blocks. Other numbers of modes may be implemented to meet the criteria of a particular application.

The motion estimation processor circuit 164 generally comprises a circuit (or module) 180 and a circuit (or module) 182. The circuit 180 may be implemented as a memory sub-system circuit. The memory sub-system circuit 180 may be configured to communicate with the external memory 162.

The circuit 182 may be implemented as a pel search circuit. The pel search circuit 182 may communicate with the external memory circuit 162 through the memory sub-system circuit 180 to receive the current blocks on which the motion estimation is to be performed. A signal (e.g., ORIG_PIXEL) may transfer the current blocks. The pel search circuit 182 may also receive reference samples stored in the external memory circuit 162 via the memory sub-system circuit 180. A signal (e.g. REF_PIXEL) may transfer the reference samples. A signal (e.g., EXT_ADDR) may be

generated by the pel search circuit 182 to read the current and reference information from the external memory circuit 162.

The signal REF_PIXEL may transfer reference frames that have been previously encoded and then decoded. The frames may represent one or more pictures used for the motion estimation. Each frame generally comprises multiple reference blocks. Each reference block may be implemented as a macroblock. The reference blocks may be arranged in a single or integer-pel resolution.

The signal ORIG_PIXEL may transfer current frames to be encoded. Each current frame generally comprises multiple current blocks for which motion estimation may be performed. Each current block may be implemented as a macroblock.

The pel search circuit 182 may be operational to determine one or more motion vectors for a current block of video data. The pel search circuit 182 may perform a variable block size motion estimation at an integer-pel resolution to determine a best partition mode and associated motion vector or motion vectors. The pel search circuit 182 generally searches all integer positions for all block-sizes within a search window. Based on integer scores, the pel search circuit 182 may identify the block size that produces the minimum sum of absolute difference score. Additional

details of the pel search circuit 182 may be found in the related United States patent applications, Serial No. 10/669,930, filed September 24, 2003, Serial No. 10/682,631, filed October 9, 2003 and Serial No. 10/___,___, filed October 22, 2003, which are hereby incorporated by reference in their entirety.

The ME processor circuit 182 generally comprises a circuit (or module) 184, a circuit (or module) 186, a circuit (or module) 188, a memory circuit (or module) 190, a circuit (or module) 192, a circuit (or module) 194, a circuit (or module) 196, a circuit (or module) 198, a circuit (or module) 200, a circuit (or module) 202, a circuit (or module) 204 and a circuit (or module) 206. The circuit 184 may receive the signal FRAME_SIG. The circuit 188 may generate the address signal EXT_ADDR. The memory circuit 190 may receive the signal REF_PIXEL. The circuit 204 may receive the signal ORIG_PIXEL. The circuit 202 may generate the signals MV, SAD and MODE.

The circuit 184 may be referred to as a start to frame circuit. The start of frame circuit 184 may be operational to determine when a new current frame begins based on the signal FRAME_SIG. The start of from circuit 184 may communicate a detected start of frame to the circuits 186 and 188.

The circuit 186 may be referred to as a reference memory coordinate calculation circuit. The reference memory coordinate calculation circuit 186 may be operational to direct reads from the external memory circuit 162 based on information received from the start of frame circuit 184. The reference memory coordinate calculation circuit 186 may calculate an X and a Y position of a macroblock column to be fetched from the external reference memory 162. Calculations of the X and the Y coordinates may be based on a global offset (gx, gy) and search width. For every X coordinate, multiple (e.g., three) Y coordinates may be calculated, one for each reference block in the macroblock column. For a macroblock row, the X coordinate generally changes every macroblock cycle and the Y coordinate change upon completion of a macroblock row.

The circuit 188 may be referred to as an external read control circuit. The external read control circuit 188 generally generates the address in the signal EXT_ADDR and associated control signals to read the macroblock column samples from the external memory circuit 162. The external read control circuit 188 may be configured to provide either a single address for a macroblock column or separates addresses for each block within the macroblock column.

The memory circuit 190 may be implemented as an internal search (or reference) memory circuit. The internal search memory circuit 190 may be operational to store reference samples copied from the external memory circuit 162 for use in the motion estimation process. The reference samples may be received by the internal search memory circuit 190 through the signal REF_PIXEL at a write port 191. The reference samples may be provided to the circuit 196 in the signal COL_SAMPLE at a read port 193. In one embodiment, the internal search memory circuit 190 may be sized to store up to twenty-four blocks of samples arranged as eight blocks horizontal by three blocks vertical. Each block may be arranged as a 16x16 array of samples. Each sample may be represented by a byte of data. The internal search memory circuit 190 may be implemented in other sizes to meet the criteria of a particular application.

The circuit 192 may be referred to as an internal write control circuit. The internal write control circuit 192 may be operational to write macroblock columns presented by the external memory circuit 162 to the internal search memory circuit 190. The internal write control circuit 192 may present a write address signal (e.g., INT_ADDR_W) to the internal search memory circuit 190

to control storage of the reference samples in the signal REF_PIXEL.

The circuit 194 may be referred to as a internal read control circuit. The internal read control circuit 194 may be operational to generate a read address signal (e.g., INT_ADDR_R). The internal read control circuit 194 may sequence reads from the internal search memory 190 to generate a signal (e.g., COL_SAMPLE). The signal COL_SAMPLE may be implemented as a sequence of columns of reference samples. The signal COL_SAMPLE may be responsive to the address signal INT_ADDR_R. In one embodiment, the signal COL_SAMPLE may transfer forty-eight reference samples in parallel simultaneously.

The circuit 196 may be referred to as a shifter circuit. The shifter circuit 196 may be operational to align the reference samples received via the signal COL_SAMPLE with appropriate inputs to the circuit 198. The shifter circuit 196 may also generate pad samples to represent non-existing reference samples outside a reference frame boundary.

The circuit 198 may be referred to as a processor circuit. The processor circuit 198 may be operational to compare an array of current samples from the current block against an array

of reference/pad samples to generate multiple scores substantially simultaneously, one score per current/reference sample comparison.

The circuit 200 may be referred to as an adder tree circuit. The adder tree circuit 200 may be operational to combine the multiple scores into an overall score for the comparison. Each overall score may be provided to the circuit 202.

The circuit 202 may be referred to as a mode decision tree circuit. The mode decision tree circuit 202 may be operational to determine a best score (e.g., sum of absolute difference) for the current block relative to the reference samples in the search window 106 loaded into the internal search memory circuit 190. The mode decision tree circuit 202 may also determine a best mode and one or more motion vectors associated with the best overall score.

The circuit 204 may be referred to as an original data control circuit. The original data control circuit 204 may be operational to control writes of current (original) samples into the processor circuit 198. The original data control circuit 204 may perform a re-order of the current samples from the current block prior to loading into the processor circuit 198.

The circuit 206 may be referred to as a datapath control circuit. The datapath control circuit 206 may be operational to control flow of the samples through the ME processor circuit 164. The datapath control circuit 206 may provide control signals to the internal read control circuit 194, the shifter circuit 196, the processor circuit 198, the adder tree circuit 200 and the mode decision tree circuit 202 and the original data control circuit 204.

Referring to FIG. 5, a partial block diagram of a second example implementation of an apparatus 210 is shown. The apparatus (or system) 210 generally comprises the external memory circuit 162 and a motion estimation processor circuit 212. The motion estimation processor circuit 212 generally comprises the memory sub-system circuit 180 and multiple pel search circuits 182a-182d. Each of the pel search circuit 182a-182d may communicate with the external memory circuit 162 through the memory sub-system circuit 180.

During motion estimation for predicted frames (e.g., P-frames), the apparatus 210 may use the pel search circuit 182a using a first global offset and the pel search circuit 182b using a second global offset. The first global offset may define a first

search window spatially in the same location in the reference frame or a location near to the current block. The first search window may determine a first motion vector generally associated with background or slow moving data. The second global offset may define a second search window spatially distant from the current block. The second search window may be determine a second motion vector generally associated with fast moving data. The pel search circuits 182a and 182b may operate together to search for a single current block at the same time, but in two different areas of the reference frame.

During a motion estimation for a bidirectional frame (e.g., B-frame), the apparatus 210 may use the pel search circuits 182a and 182b for forward prediction. The pel search circuits 182c and 182d may be used for backwards prediction. Other arrangements and uses of the pel search circuits 182a-182d may be implemented to meet the criteria of a particular application. Therefore, the apparatus 210 may search over what appears to be a large search window without (i) consuming the bandwidth of the external memory circuit 162 or (ii) the storage capacity of the internal search memory circuit 190 that would be used to actually copying the large search window.

The apparatus 100 and the apparatus 210 may each provide two small independent search windows to capture background/small motion and fast motion for high picture quality. The boxcar search window approach of the present invention generally minimizes memory bandwidth for the external memory circuit 162 and may keep data flows regular. Therefore, a complexity of the control hardware may be kept low. The global offsets may be adjustable on a per frame basis. Adjusting the global offsets generally permits the boxcar search windows to adapt to fast motions (i) within a frame of occurrence and (ii) immediately if a sub-sampled search of a current frame is used. The present invention may be adapted for use in encoders with motion estimation processors for any other digital video compression process (e.g., proprietary, Windows Media Video 9 series (Microsoft Corp., Redmond, Washington), DivX® (Project Mayo, San Diego, California) , MPEG-1/2/4 (Moving Pictures Expert Group, International Organization for Standards, Geneva, Switzerland), H.261/3/4 (International Telecommunication Union Telecommunication Standardization Sector, Geneva, Switzerland)).

As used herein, the term "simultaneously" is meant to describe events that share some common time period but the term is

not meant to be limited to events that begin at the same point in time, end at the same point in time, or have the same duration.

While the invention has been particularly shown and described with reference to the preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made without departing from the spirit and scope of the invention.